The Bell System Technical Journal
Vol. 50, No. 6, July-August, 1971
Printed in U.S.A.

Some Analytical Investigations of Phase Contrast Imaging

By C. M. NAGEL, JR.

(Manuscript received December 1, 1970)

An analytical study of phase contrast imaging is performed. It is shown that the complex disturbance in the image plane can be represented as a convolution of the object disturbance and the Fourier transform of the transmission function of the phase plate. The simple theory of the phase contrast microscope is then derived as a limiting case of this more general result that is applicable when the size of the phase object is small compared to the area of the entrance aperture of the system. The response of a general system to several simple large phase objects is also examined, and it is shown that qualitative information about these objects can be obtained from the intensity pattern when the phase perturbations are small, providing the background is sufficiently uniform, and the size of the phase spot on the phase plate is carefully chosen. The study provides insight into the type of performance that can be achieved, for example, if phase contrast imaging is used to extract phase information from an optical memory or if it is used as an experimental tool to study the qualitative behavior of such phenomena as clear air turbulence.

1. INTRODUCTION

One of the most popular techniques for converting a spatial phase variation into a spatial intensity variation is the phase contrast imaging scheme originally proposed by F. Zernike[1-3] in 1935. Proceeding mainly from a heuristic point of view, Zernike recognized that, with a conventional imaging system, most of the direct light passes through a small region R in the focal plane while, in many instances, most of the diffracted light is scattered away from this region. If a plate whose transmission function is given by

$$T = \begin{cases} ae^{i\alpha} & \text{in } R \\ 1 & \text{otherwise} \end{cases} \tag{1}$$

is placed in the focal plane, the direct light can be modified and made to interfere with the diffracted light to produce an intensity pattern in the image plane. Moreover, if the phase perturbations ϕ that are introduced by the object are small enough so that the approximation

$$e^{j\phi} \approx 1 + j\phi$$

holds, the intensity in the image plane and the phase of the object may be linearly related. These concepts led to the design of the phase contrast microscope.

The key assumption in the theory is that the effect of the phase plate on the diffracted light is negligible. In 1953, H. H. Hopkins[4] rigorously demonstrated that, if this assumption holds, a phase object will produce an intensity pattern given by

$$I = \frac{1}{M^2} \mid ae^{j\alpha} - 1 + e^{j\phi} \mid^2. \tag{2}$$

In this expression, M is the magnification of the system, and it is assumed in the derivation that the object is illuminated by a unit amplitude monochromatic plane wave. When ϕ is small, a linear relationship,

$$I = \frac{1}{M^2} (a^2 + 2a\phi \sin \alpha), \tag{3}$$

results, as predicted by Zernike.†

There are many potential applications of phase contrast imaging in which the basic assumption employed in Refs. 1–5 will not be valid. Therefore, it is interesting to examine what results will follow if these assumptions are not made. In this regard, Hopkins[6] rigorously studied the diffraction images of circular disks under coherent illumination, and M. De and S. C. Som,[7] and De and P. K. Mondal,[8] investigated the images produced by similar objects under incoherent illumination. In this paper, we will derive the theory from the point of view of Fourier optics and demonstrate that, in general, the complex disturbance over the image plane can be represented by a convolution of the object disturbance and the Fourier transform of the transmission function of the phase plate. The approximate theory, equations (2) and (3), is then shown to be a limiting case of the general theory, applicable if the size of the phase object is small compared to the entrance aperture of the system. Such is usually the case when the

† A related analysis is also given in Born and Wolf,[5] Chapter 8.

magnification of the system is large. For larger objects, the convolution integral yields a complicated relationship between the phase of the object and the intensity of the image.

The response of the system to several simple large phase objects is then studied, and it is shown that qualitative information about the phase distributions of these objects can be obtained when the phase perturbations are small if an appropriate phase plate is employed.

The work that is discussed in this paper complements the experimental investigations carried on at Bell Telephone Laboratories by N. J. Kolettis.

II. GENERAL THEORY

The optical system that we will consider is shown in Fig. 1. A circular entrance aperture with radius R_o is located a distance d_o in front of the lens, and a phase plate with a transmission function

$$T_f(r_f) = \begin{cases} ae^{ja} & r_f \leqq R_f^o \\ 1 & R_f^o < r_f \leqq R_{fo} \\ 0 & R_{fo} < r_f \end{cases} \tag{4}$$

is centered in the focal plane. The phase object is assumed to occupy some portion of the entrance aperture, and the illuminating light is taken to be a normally incident unit amplitude monochromatic plane wave.

Let the complex disturbance in the object plane be denoted by $U_o(x_o, y_o)$ where U_o is assumed to be zero at points outside the entrance aperture. Then, from the Fresnel approximation,[9] it follows

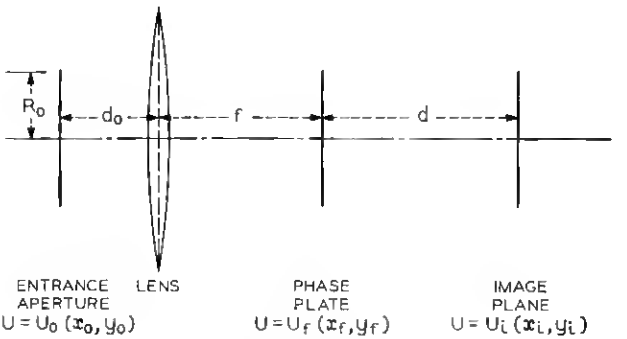

Fig. 1—System geometry.

that the complex disturbance in the focal plane to the left of the phase plate is

$$U_f(x_f , y_f) = \frac{\exp [jk(d_o + f + n\,\Delta_o)]}{j\lambda f} \exp\left[\frac{jk}{2f}\left(1 - \frac{d_o}{f}\right)(x_f^2 + y_f^2)\right]$$
$$\cdot \iint_{-\infty}^{\infty} U_o(x_o , y_o) P_L\left(x_o + \frac{d_o}{f} x_f , y_o + \frac{d_o}{f} y_f\right)$$
$$\cdot \exp\left[\frac{-jk}{f}(x_o x_f + y_o y_f)\right] dx_o\, dy_o . \quad (5)$$

In this expression f is the focal length of the lens, λ is the wavelength of the illuminating light, k is the wavenumber $2\pi/\lambda$, and $kn\Delta_o$ is the phase shift introduced by the lens. P_L is the pupil function of the lens which assumes the value of unity over the region covered by the lens, and is zero elsewhere.

Now, the complex disturbance on the right side of the phase plate is simply $U_f T_f$, and, therefore, if the Fresnel approximation is also used to describe the propagation from this plane to the image plane, the complex disturbance in the image plane becomes

$$U_i(x_i , y_i)$$
$$= \frac{K}{\lambda^2 f d} \iint_{-\infty}^{\infty} \iint_{-\infty}^{\infty} U_o(x_o , y_o) P_L\left(x_o + \frac{d_o}{f} x_f , y_o + \frac{d_o}{f} y_f\right) T_f(x_f , y_f)$$
$$\cdot \exp\left\{-\frac{j2\pi}{\lambda d}[(Mx_o + x_i)x_f + (My_o + y_i)y_f]\right\} dx_o\, dy_o\, dx_f\, dy_f , \quad (6)$$

where

$$K = -\exp [jk(d_o + f + d + n\,\Delta_o)] \exp\left[\frac{jk}{2d}(x_i^2 + y_i^2)\right].$$

Here we are assuming that the distances d_o, d, and f obey the lens law,

$$\frac{1}{f}\left(1 - \frac{d_o}{f}\right) + \frac{1}{d} = 0,$$

and we have denoted the magnification of the system, d/f, by M.

The integral in (6) is complicated by the coupling of the coordinates x_o, x_f and y_o, y_f via the pupil function of the lens. This phenomenon, known as the vignetting effect,[9] can be ignored if it is assumed that the lens is large enough so that the sum of the squares of the arguments of P_L is always less than the square of the radius of the lens

for all values of x_o, x_f, y_o, and y_f for which U_o and T_f are nonzero.† With this assumption P_L may be taken as one.

To further simplify equation (6) we introduce the notation

$$\tilde{x}_i = \frac{x_i}{M}, \qquad \tilde{y}_i = \frac{y_i}{M}, \qquad x = x_o + \tilde{x}_i, \qquad y = y_o + \tilde{y}_i.$$

Then, neglecting the vignetting effect,

$$U_i(x_i, y_i) = \frac{K}{\lambda^2 fd} \iint_{-\infty}^{\infty} U_o(x - \tilde{x}_i, y - \tilde{y}_i) F[T_f]\, dx\, dy$$

where

$$F[T_f] = \iint_{-\infty}^{\infty} T_f(x_f, y_f) \exp\left[-\frac{j2\pi M}{\lambda d}(xx_f + yy_f)\right] dx_f\, dy_f. \quad (7)$$

We observe that the disturbance in the image plane can be written as a convolution of the disturbance in the object plane and the Fourier transform of the transmission function of the phase plate evaluated at the spatial frequencies $Mx/\lambda d$, $My/\lambda d$.

Let us now examine the form of $F[T_f]$ more carefully. The transmission function of the phase plate, from (4), is given by

$$T_f(r_f) = (ae^{j\alpha} - 1)\ \mathrm{circ}\ (r_f/R_f^o) + \mathrm{circ}\ (r_f/R_{fo})$$

with

$$r_f = (x_f^2 + y_f^2)^{\frac{1}{2}}$$

and

$$\mathrm{circ}\ (r_f/R) = \begin{cases} 1 & r_f \leqq R \\ 0 & r_f > R. \end{cases}$$

It follows then that

$$F[T_f] = \frac{\lambda^2 d^2}{M^2}\left[(ae^{j\alpha} - 1)\frac{MR_f^o}{\lambda d}\frac{J_1\left(\frac{2\pi MR_f^o}{\lambda d} r\right)}{r} + \frac{MR_{fo}}{\lambda d}\frac{J_1\left(\frac{2\pi MR_{fo}}{\lambda d} r\right)}{r}\right] \quad (8)$$

where

$$r = (x^2 + y^2)^{\frac{1}{2}},$$

and J_1 is the Bessel function of the first kind of order one.

† Recall $U_o = 0$ for $x_o^2 + y_o^2 > R_o$, and $T_f = 0$ for $x_f^2 + y_f^2 > R_{fo}$.

Now, at optical wavelengths, R_{f_o}/λ is exceedingly large, and for all practical purposes, the second term in (8) can be approximated by a two-dimensional delta function with little error.† On the other hand, R_f^o is typically of the order of the radius of the Airy disk of the diffraction pattern of the entrance aperture so that the first term cannot be so simplified. To emphasize this fact, R_f^o will be written as

$$R_f^o = \gamma \frac{\lambda f}{R_o}$$

where γ is a constant of the order of unity. ($\gamma \approx 0.61$ if R_f^o is the radius of the Airy disk.)

With this notation and the definition of M, equation (8) becomes

$$F[T_f] \approx \frac{\lambda^2 d^2}{M^2}\left[(ae^{j\alpha} - 1)\frac{\gamma}{R_o}\frac{J_1\left(\frac{2\pi\gamma}{R_o}r\right)}{r} + \delta(r)\right]. \tag{9}$$

Returning to the expression for the complex disturbance in the image plane (7), we conclude that

$$U_i(x_i, y_i) \approx \frac{K}{M}\left[(ae^{j\alpha} - 1)\frac{\gamma}{R_o}\iint_{-\infty}^{\infty} U_o\left(x - \frac{x_i}{M}, y - \frac{y_i}{M}\right) \cdot \frac{J_1\left[\frac{2\pi\gamma}{R_o}(x^2 + y^2)^{\frac{1}{2}}\right]}{(x^2 + y^2)^{\frac{1}{2}}} dx\, dy + U_o\left(-\frac{x_i}{M}, -\frac{y_i}{M}\right)\right]. \tag{10}$$

Equation (10) is the fundamental result that will be employed throughout the remainder of this paper.

III. SMALL PHASE OBJECTS–THE PHASE CONTRAST MICROSCOPE

The form of equation (10) indicates that, in general, a complicated relationship exists between the object and the image. We will now demonstrate, however, that the simple theory, equations (2) and (3), may be employed *if the ratio of the area of the phase object to the area of the entrance aperture is sufficiently small.* This is usually the case for a microscopic system.

We will assume that the phase object is centered in the entrance

† We shall observe later that this approximation yields sharp jumps in intensity when the phase object possesses phase discontinuities. In practice some smoothing of the discontinuities in the intensity pattern will always result. (See Ref. 10 for example.) We are neglecting these lower order effects for simplicity.

aperture and write $U_o(x_o , y_o)$ as

$$U_o(x_o , y_o) = \text{circ}\left[\frac{(x_o^2 + y_o^2)^{\frac{1}{2}}}{R_o}\right] + (e^{j\phi} - 1)P_o ,$$

where P_o has a value of one over the region covered by the phase object and a value of zero elsewhere. In the limit of a very small phase object,† the contribution of the second term in U_o to the *integral* in (10) approaches zero, and the integral can be approximated by

$$\mathrm{I} = \frac{\gamma}{R_o}\int\int_{-\infty}^{\infty} \text{circ}\left[\frac{((x - \tilde{x}_i)^2 + (y - \tilde{y}_i)^2)^{\frac{1}{2}}}{R_o}\right] \frac{J_1\left[\frac{2\pi\gamma}{R_o}(x^2 + y^2)^{\frac{1}{2}}\right]}{(x^2 + y^2)^{\frac{1}{2}}}\, dx\, dy.$$

The convolution is most easily evaluated by first taking the Fourier transform of I and then taking the inverse transform of the results. These two operations yield

$$\mathrm{I} = \int_0^{2\pi\gamma} J_1(t) J_0\left(\frac{r_i}{MR_o}\, t\right) dt \tag{11}$$

where r_i is the radial coordinate in the image plane.

If we now restrict the coordinates in the image plane to the region corresponding to the image of the phase object, the ratio r_i/MR_o will be small, and I can be approximated by

$$\mathrm{I} \approx \int_0^{2\pi\gamma} J_1(t)\, dt = 1 - J_0(2\pi\gamma).$$

Finally, with an appropriate choice of the radius of the central spot of the phase plate (and thus γ), the contribution from the Bessel function $J_0(2\pi\gamma)$ can be made to vanish, and we conclude that, for a sufficiently small phase object and a suitable γ,

$$U_i(x_i , y_i) \approx \frac{K}{M}\left[(ae^{j\alpha} - 1) + \exp\left[j\phi\left(-\frac{x_i}{M} , -\frac{y_i}{M}\right)\right]\right].$$

The intensity of the image is then

$$I(x_i , y_i) \equiv |\, U_i \,|^2 \approx \frac{1}{M^2}\left|\, (ae^{j\alpha} - 1) + \exp\left[j\phi\left(-\frac{x_i}{M} , -\frac{y_i}{M}\right)\right] \right|^2$$

which agrees with equation (3). The simple theory of the phase contrast microscope is, therefore, a limiting case of the more general approach.

† Consider a single cell under a microscope, for example.

IV. MACROSCOPIC SYSTEMS

In many potential applications of phase contrast imaging the magnification of the system may be of the order of unity, and the area of the phase object may be comparable to the area of the entrance aperture. In these cases, the approximations that were employed in the last section cannot be used, and the complex disturbance in the image plane can, in general, only be determined via a direct application of equation (10). In order to provide some representative results, we have chosen a few simple, yet important, examples for which the integral in (10) can be easily evaluated.

4.1 *Background*

The purpose of any phase contrast system is to give a visual representation of the phase distribution in the entrance aperture. If a system is to yield meaningful results, therefore, the intensity of the image of the aperture should be essentially constant when no phase objects are present. In this case

$$U_o(x_o\ ,\ y_o) = \text{circ}\left[\frac{(x_o^2 + y_o^2)^{\frac{1}{2}}}{R_o}\right]$$

and from (10) and (11) it follows that

$$I(r_i) = \frac{1}{M^2}\left|(ae^{j\alpha} - 1)\int_0^z J_1(t)J_0\left(\frac{r_i}{MR_o}t\right)dt + 1\right|^2 \qquad (r_i \leqq MR_o) \tag{12}$$

where the notation $z = 2\pi\gamma$ has been introduced.

For our later purposes, it is convenient to define a normalized intensity, $\bar{I} = M^2 I$. Plots of this function are shown in Figs. 2–5 for various values of z. Clearly, the response of the system to the illuminating light is critically dependent upon the size of the central spot of the phase plate. Where $a = 1$ and $\alpha = \pi/2$, a typical case that is employed in phase contrast imaging, near uniform backgrounds result for values of z corresponding to spot sizes on the order of one-half the Airy disk or smaller.† For larger z, tbe background varies considerably until several rings in the diffraction pattern of the entrance aperture are covered by the phase spot.‡ We shall see later that these variations produce distortions in the intensity pattern of phase objects when tbey

† The task of making a phase plate with a spot size this small may be difficult.
‡ The phase contrast system will begin to behave like a direct imaging system for spot sizes larger than those shown, and thus large z values are not practical.

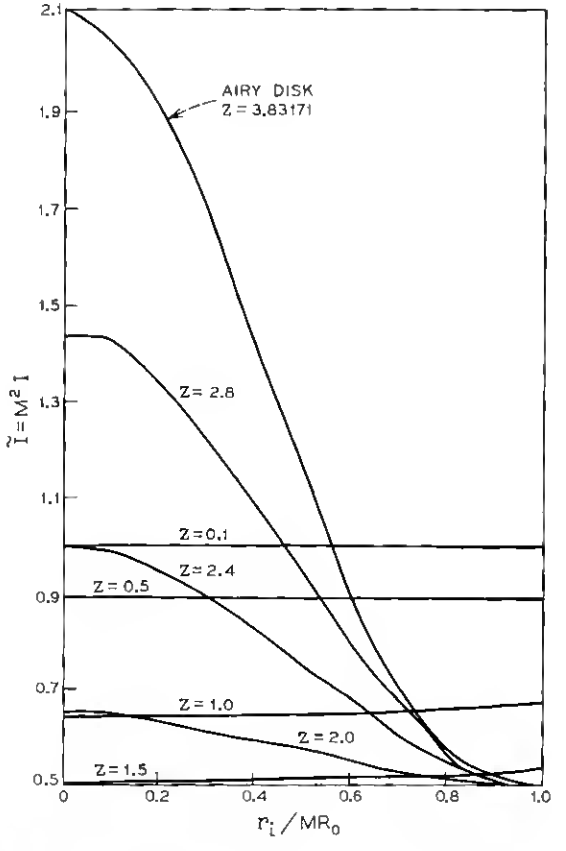

Fig. 2—Background intensity distributions for spot sizes less than or equal to the Airy disk ($a = 1$, $\alpha = \pi/2$).

are placed in the aperture. The case $a = 1$, $\alpha = \pi/2$, $z = 1.5$ will be of special interest to us, since in this case the integral in (12) is approximately 1/2 except when r_i is near the edge of the image of the aperture, the background is essentially constant, and several interesting results can be obtained.

Frequently, a phase contrast system will employ a partially absorbing phase plate ($a < 1$) to improve the contrast between the image of the phase object and the background [see equation (3)]. Background intensities for $a = 0.1$, $\alpha = \pi/2$ and various values of z are shown in Figs. 4 and 5. There it may be observed that, in general, the background intensity is reduced, although a bright ring appears in the

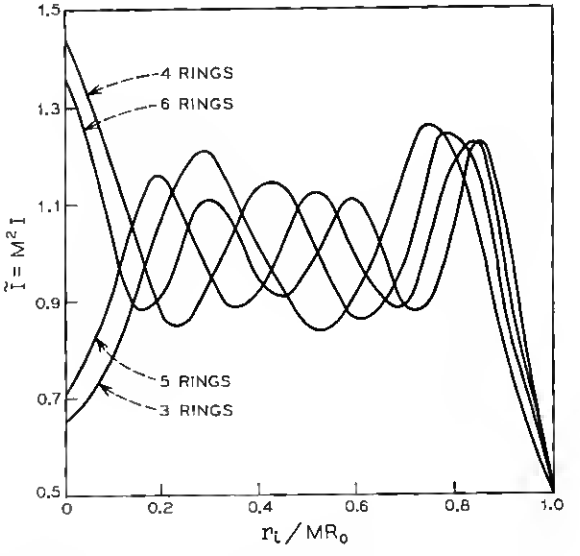

Fig. 3—Background intensity distributions for spot sizes larger than the Airy disk ($a = 1$, $\alpha = \pi/2$).

region close to the edge of the image of the aperture. (This ring has been observed experimentally by Kolettis.) If this edge effect is ignored, near uniform backgrounds are obtained for values of z that include five or six Airy rings.

Taken together, the above results indicate that care must be exercised in designing a phase contrast system if near uniform backgrounds are to be achieved.

4.2 *Phase Disk*

Consider the "phase disk" defined by

$$U_o(x_o, y_o) = (e^{i\phi} - 1)\ \text{circ}\left[\frac{(x_o^2 + y_o^2)^{\frac{1}{2}}}{R_o^{(1)}}\right] + \text{circ}\left[\frac{(x_o^2 + y_o^2)^{\frac{1}{2}}}{R_o}\right]$$

$$(0 \leqq R_o^{(1)} \leqq R_o) \qquad (13)$$

where ϕ is a constant. The geometry corresponds to a disk with radius $R_o^{(1)}$ centered in the entrance aperture which introduces a constant phase shift ϕ in the incident wave. Substituting in equation (10) and using the technique that was employed to derive equation (11), we obtain the intensity distribution

$$\tilde{I}(r_i) = \begin{cases} |\,(ae^{j\alpha} - 1)[(e^{j\phi} - 1)\mathrm{I}^{(1)}(r_i) + \mathrm{I}(r_i)] + e^{j\phi}\,|^2, \\ \qquad\qquad r_i \leqq MR_o^{(1)} \\ |\,(ae^{j\alpha} - 1)[(e^{j\phi} - 1)\mathrm{I}^{(1)}(r_i) + \mathrm{I}(r_i)] + 1\,|^2, \\ \qquad\qquad MR_o^{(1)} < r_i \leqq MR_o\,. \end{cases} \tag{14}$$

In this expression $\mathrm{I}(r_i)$ is the integral given in (11) and

$$\mathrm{I}^{(1)}(r_i) = \frac{R_o^{(1)}}{R_o} \int_0^{2\pi\gamma} J_1\left(\frac{R_o^{(1)} t}{R_o}\right) J_0\left(\frac{r_i}{MR_o}\, t\right) dt. \tag{15}$$

Note that the integral I only takes into account the finite extent of the entrance aperture, while $\mathrm{I}^{(1)}$ includes the size of the phase object as well.

Plots of (14) are shown in Figs. 6–8 for $a = 1$, $\alpha = \pi/2$ and $z = 1.5$. The value of z has been chosen to correspond to a near uniform back-

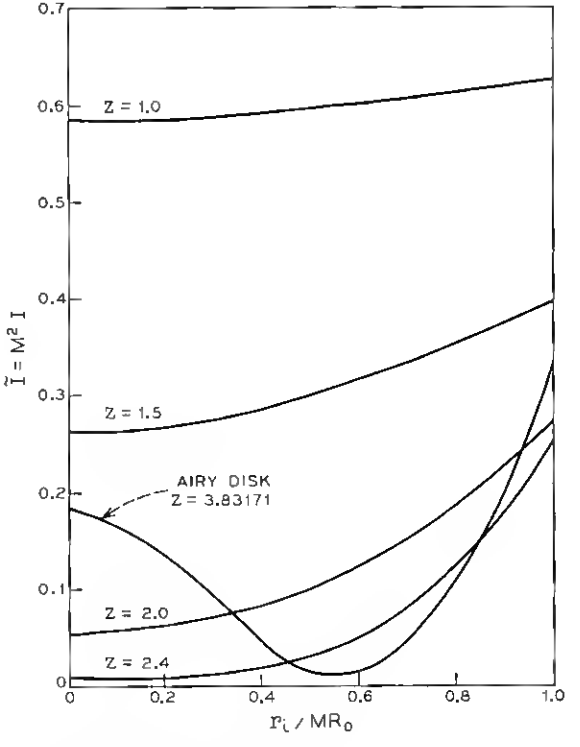

Fig. 4—Background intensity distributions for spot sizes less than or equal to the Airy disk ($a = 0.1$, $\alpha = \pi/2$).

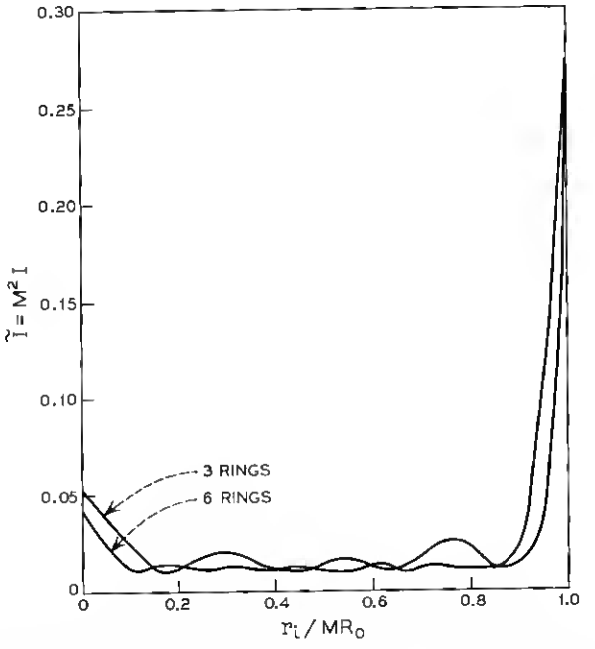

Fig. 5—Background intensity distributions for spot sizes larger than the Airy disk ($a = 0.1$, $\alpha = \pi/2$).

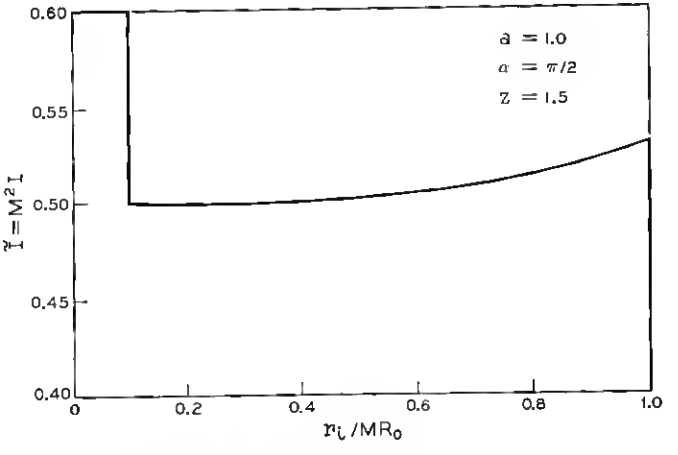

Fig. 6—Response to a phase disk ($\phi = 0.1$, $R_s^{(1)}/R_s = 0.1$).

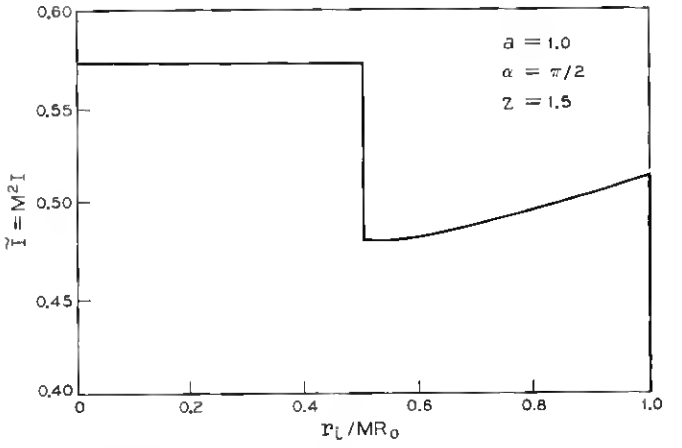

Fig. 7—Response to a phase disk ($\phi = 0.1, R_o^{(1)}/R_o = 0.5$).

ground when $R_o^{(1)} = R_o$ or when $R_o^{(1)} = 0$. It is clear from the figures that, for the small value of ϕ chosen, the qualitative behavior of the phase distribution in the entrance aperture is reproduced quite nicely. The phase disk can be readily observed and the magnitude of the discontinuity in $\tilde{I}$ is on the order of ϕ.†

In Fig. 9 we have indicated the response of the same system with the exception that ϕ has now been assigned a value of π. Note that in this case, the jump in $\tilde{I}$ depends critically upon the size of the phase object. Moreover, it appears that a value of $R_o^{(1)}$ exists for which no contrast results and beyond which the contrast is reversed. This suggests that for larger ϕ values, the system performs quite poorly.

To explain this behavior, we need only to return to equation (14) from which it immediately follows that the magnitude of the discontinuity in $\tilde{I}$ is

$$\Delta\tilde{I} = 2 \operatorname{Re}\{(ae^{j\alpha} - 1)(e^{j\phi} - 1)[(e^{j\phi} - 1)\mathrm{I}^{(1)}(MR_o^{(1)}) + \mathrm{I}(MR_o^{(1)})]\}$$
$$= 2(a\cos\alpha - 1)(1 - \cos\phi)(2\mathrm{I}^{(1)} - \mathrm{I}) + 2a\mathrm{I}\sin\alpha\sin\phi. \quad (16)$$

For small ϕ,

$$\Delta\tilde{I} \approx (2a\mathrm{I}\sin\alpha)\phi,$$

while for larger ϕ, the full equation (16) must be employed. Recalling, that for $z = 1.5$, $\mathrm{I} \approx 0.5$ for most values of r_i, we conclude that, for

† The increase in intensity in the outer portion of the image is due to the slight nonuniformity of the background in this region.

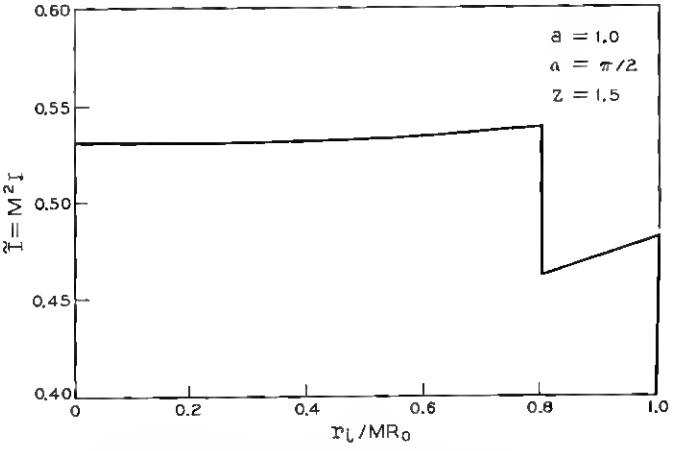

Fig. 8—Response to a phase disk ($\phi = 0.1$, $R_0^{(1)}/R_0 = 0.8$).

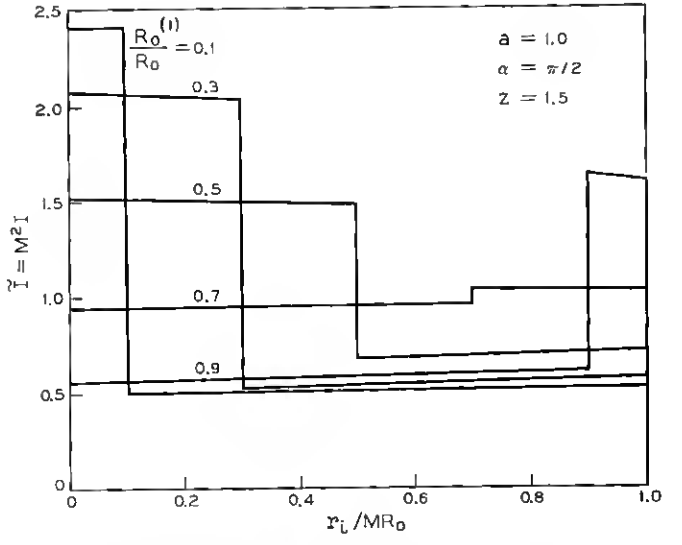

Fig. 9—Response to phase disks of various sizes ($\phi = \pi$).

$a = 1$, $\alpha = \pi/2$, $z = 1.5$, and ϕ small,

$$\Delta\tilde{I} \approx \phi$$

as observed.

When $\phi = \pi$, however, equation (16) reduces to

$$\Delta\tilde{I} = 4(a \cos \alpha - 1)[2\mathrm{I}^{(1)}(MR_o^{(1)}) - \mathrm{I}(MR_o^{(1)})].$$

From the definitions of the integrals (11) and (15), it is clear that for $\phi = \pi$ there exists a value of $R_o^{(1)}$ for which $\Delta\tilde{I} = 0$ and beyond which the contrast is reversed.

In the preceding section, we observed that nonuniform backgrounds can result when the size of the central spot of the phase plate is not carefully chosen. Figure 10 demonstrates the response of a phase contrast system to a phase disk when this is the case. The value of z has been chosen so that the spot size corresponds to the Airy disk of the diffraction pattern of the entrance aperture. The background corresponding to this value of z is shown in Fig. 2. Note that the phase disk simply perturbs the background pattern yielding an intensity profile that does not have the desired step function shape.

4.3 *Phase Rings*

The analytical results presented in Section 4.2 indicate that a discontinuity in the radial coordinate of the phase distribution of the entrance aperture can be qualitatively reproduced in the intensity pattern providing a sufficiently uniform background is chosen and the magnitude of the discontinuity is small. In other cases, such as a large discontinuity or a nonuniform background, rather poor results are obtained. We will now generalize these results to a more complicated radial phase distribution.

Suppose the phase distribution in the entrance aperture consists of a central disk and a sequence of annular rings in each of which the phase is some constant ϕ_k. Let the radii defining the location of the phase discontinuities be $R_o^{(k)}$, $k = 1, \cdots, n-1$, and let $R_o^{(n)} = R_o$. Then U_o can be written as

$$U_o(x_o, y_o) = \sum_{k=1}^{n} U_o^{(k)} \operatorname{circ}\left[\frac{(x_o^2 + y_o^2)^{\frac{1}{2}}}{R_o^{(k)}}\right] \tag{17}$$

where

$$U_o^{(k)} = \begin{cases} \exp[j\phi_k] - \exp[j\phi_{k+1}] & k = 1, \cdots, n-1 \\ \exp[j\phi_n] & k = n. \end{cases}$$

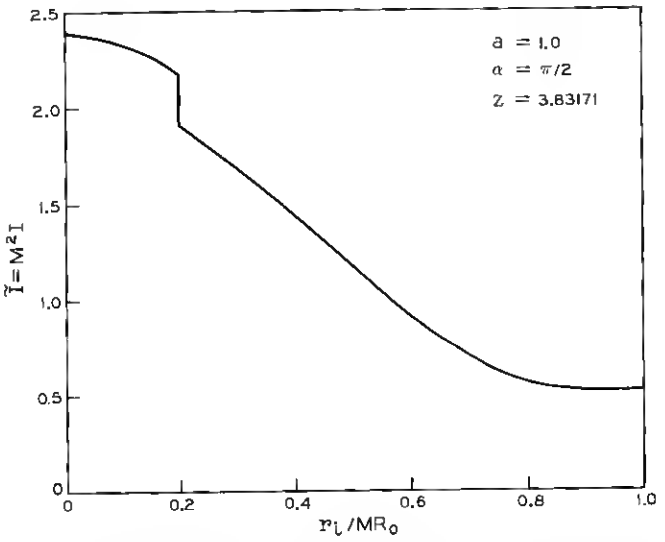

Fig. 10—Response to a phase disk when the phase plate spot size equals the Airy disk of the entrance aperture ($\phi = 0.1$, $R_o^{(1)}/R_o = 0.2$).

The normalized intensity distribution immediately follows as

$$\tilde{I}(r_i) = \begin{cases} \left| (ae^{i\alpha} - 1) \sum_{k=1}^{n} U_o^{(k)} \mathrm{I}^{(k)}(r_i) + \exp[j\phi_1] \right|^2 \\ \qquad 0 \leqq r_i \leqq MR_o^{(1)} \\ \left| (ae^{i\alpha} - 1) \sum_{k=1}^{n} U_o^{(k)} \mathrm{I}^{(k)}(r_i) + \exp[j\phi_{l+1}] \right|^2 \\ \qquad MR_o^{(l)} < r_i \leqq MR_o^{(l+1)}, \quad l = 1, \cdots, n-1. \end{cases} \tag{18}$$

In this expression $\mathrm{I}^{(k)}(r_i)$ is the generalization of (15) which is given by

$$\mathrm{I}^{(k)}(r_i) = \frac{R_o^{(k)}}{R_o} \int_0^z J_1\left(\frac{R_o^{(k)}}{R_o}\, t\right) J_0\left(\frac{r_i}{MR_o}\, t\right) dt. \tag{19}$$

Plots of (18) are shown in Figs. 11–13 for $a = 1$, $\alpha = \pi/2$, and $z = 1.5$. The assumed phase distributions in the entrance aperture are given at the top of the figures and the resulting intensity profiles are shown below. A qualitative reproduction of the phase distributions results in each case.

The behavior of (18) in the limit of small phases, ϕ_k, explains some of the success that was achieved in the figures. From the definition of the

coefficients $U_o^{(k)}$ it follows that, if all the ϕ_k are small and, in particular, if $\phi_n = 0$,

$$U_o^{(k)} \approx \begin{cases} j(\phi_k - \phi_{k+1}) & k = 1, \cdots, n-1 \\ 1 & k = n. \end{cases}$$

Therefore, the jumps in intensity at the points $r_i = MR_o^{(k)}$, to first order in the phase jumps, is just

$$\begin{aligned} \Delta\tilde{I}(MR_o^{(k)}) &\approx 2 \text{ Re } \{(ae^{j\alpha} - 1)j(\phi_{k+1} - \phi_k)\mathrm{I}^{(n)}(MR_o^{(k)})\} \\ &= 2a \sin \alpha(\phi_k - \phi_{k+1})\mathrm{I}^{(n)}(MR_o^{(k)}). \end{aligned} \tag{20}$$

Now, in the cases shown in the figures, $a = 1$, $\alpha = \pi/2$, and $z = 1.5$. Recalling that for $z = 1.5$, $\mathrm{I}^{(n)} \approx 0.5$ we conclude that

$$\Delta\tilde{I}(MR_o^{(k)}) \approx \phi_k - \phi_{k+1}.$$

Thus, if the ϕ_k are sufficiently small, and an appropriate value of z

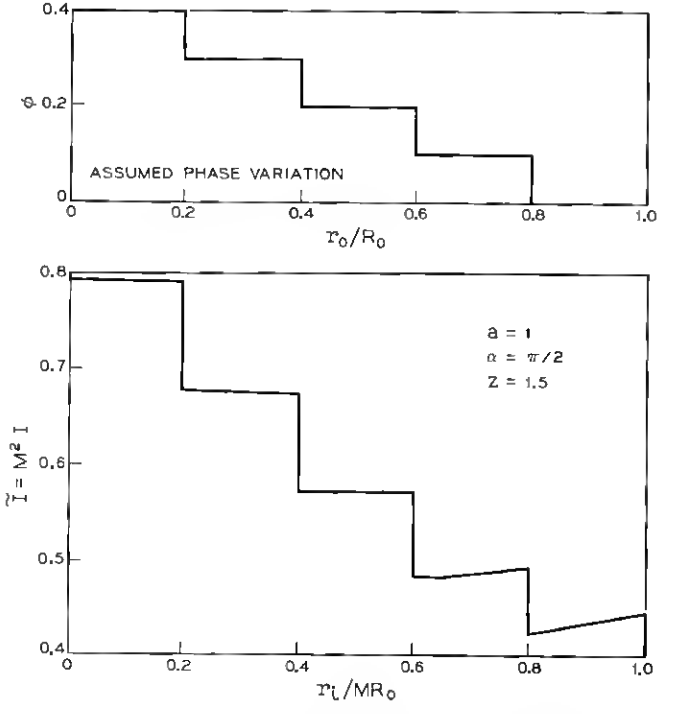

Fig. 11—Response to phase rings with decreasing phase.

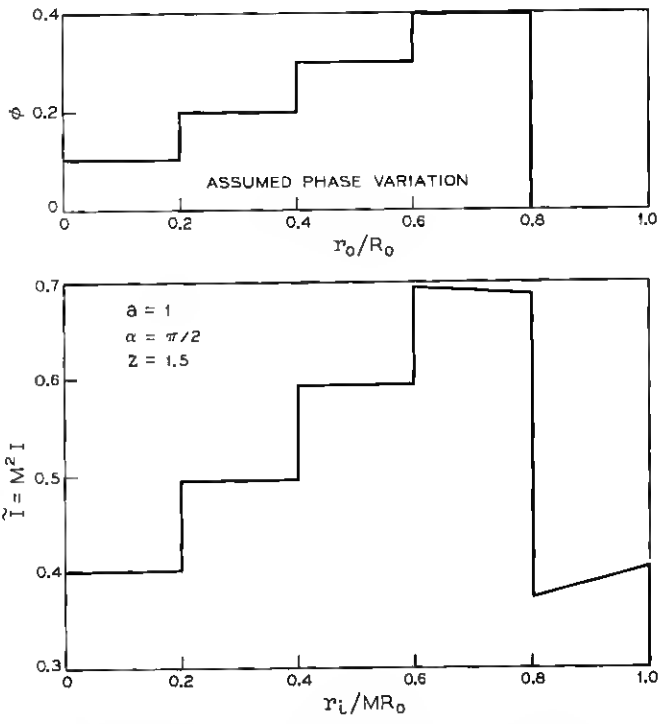

Fig. 12—Response to phase rings with increasing phase.

is chosen, the jumps in intensity are of the order of the jumps in phase.

Figures 11–13 also indicate that, for the values of the parameters chosen, the piecewise constant behavior of the phase distributions is essentially reproduced in the intensity profile. The success here is mainly due to the fact that not only was z chosen to yield a nearly uniform background, but also the z value selected was relatively small. Under these conditions the dependence of the integrals $\mathbf{I}^{(k)}(r_i)$ is suppressed,† and the effect of the phase object is to introduce essentially piecewise constant perturbations to a nearly uniform background.

To demonstrate what effects can occur for a larger z value, we have plotted in Fig. 14 the response of a phase contrast system to a sequence of phase rings when $a = 0.1$, $\alpha = \pi/2$, and $z = 21.21163$.‡ This is a practical case to consider since small values of a are fre-

† Consider equation (19) in the limit of small z.

‡ This value of z corresponds to a phase plate spot size that includes about five and one-half Airy rings.

quently employed in phase constant microscopy to improve the contrast between the phase object and the background. (For our purposes, large z values are needed to give a fairly uniform background over a large portion of the image of the entrance aperture when a is small.) Unfortunately, the results that are observed are then quite poor. The

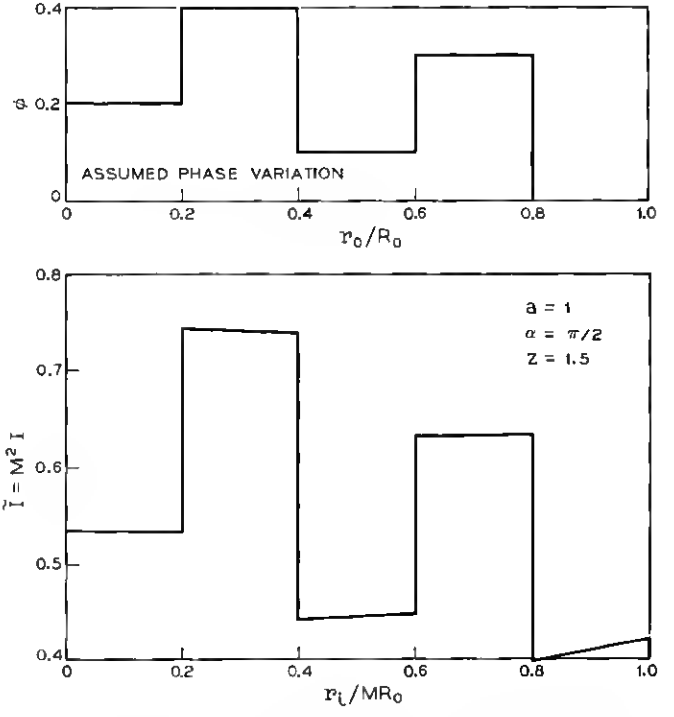

Fig. 13—Response to phase rings with alternating phase.

problem here is that the integrals $I^{(k)}(r_i)$ are now highly nonlinear functions of r_i, and, therefore, nonlinear perturbations of the order of the background are introduced.†

4.4 *Semi-circular Phase Disks*

The results of the last section demonstrate that with an appropriate choice of the parameters a, α, and z, circularly symmetric phase

† Figure 5 demonstrates that for the values of the parameters chosen, the background is of the order of $a^2 = 0.01$. However, from equation (20), we observe that the jumps in $\tilde{I}$, which we may take as indications of the magnitudes of the perturbations of the background, are of the same order when $\phi_k - \phi_{k+1}$ is small.

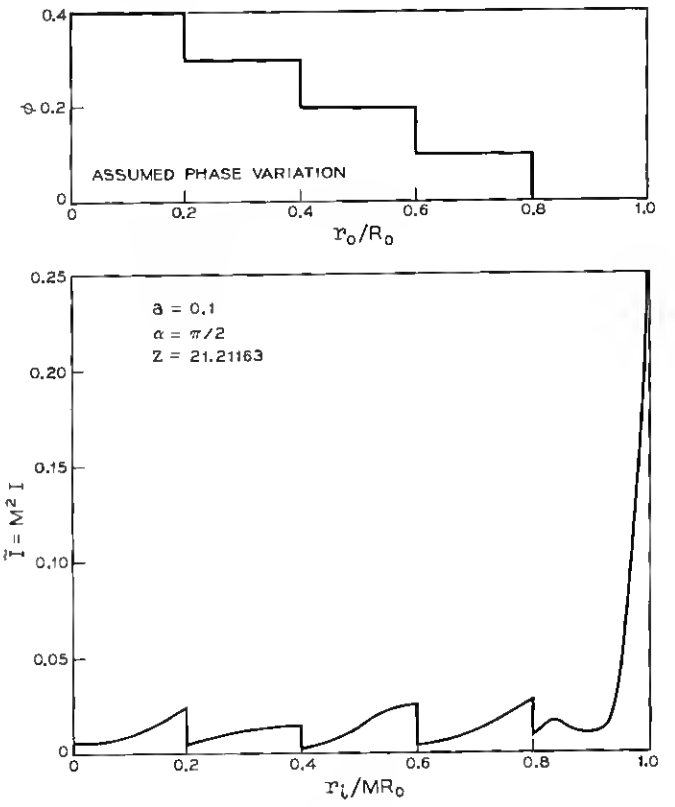

Fig. 14—Response to phase rings ($a = 0.1$, $z = 21.21163$).

distributions possessing discontinuities in the radial coordinate can be qualitatively reproduced in the intensity patterns. As a final example, we now examine the response of a phase contrast system to a simple angular discontinuity.

Let the coordinates in the object plane be r_o and θ_o and consider the complex disturbance

$$U_o(r_o, \theta_o) = \text{cire}\,(r_o/R_o) \exp [j\Theta(\theta_o)], \tag{21}$$

where

$$\Theta(\theta_o) = \begin{cases} \gamma_1 & 0 \leqq \theta_o \leqq \pi \\ \gamma_2 & \pi < \theta_o < 2\pi \end{cases}$$

with γ_1 and γ_2 assumed constant. With this phase object, the entrance aperture is divided into two semi-circular regions in each of which the phase is constant.

In order to calculate the intensity pattern that is produced by (21), we must compute the convolution integral (see equation (10)):

$$I_s = U_o * \frac{\gamma}{R_o} \frac{J_1(2\pi\gamma r/R_o)}{r}. \tag{22}$$

As in our previous analysis, we will accomplish this task by first taking the Fourier transform of I_s and then taking the inverse transform of the result.

Now, it is a well-known result in the theory of Fourier transforms[9] that if a function $g(r, \theta)$ is separable in r and θ, i.e., $g(r, \theta) = g_r(r)g_\theta(\theta)$, then the Fourier transform of g can be expressed by the following infinite series of Hankel transforms:

$$F[g] = \sum_{k=-\infty}^{\infty} c_k(-j)^k e^{jk\phi} H_k[g_r(r)]$$

where

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} g_\theta(\theta) e^{-jk\theta}\, d\theta$$

and

$$H_k[g_r(r)] = 2\pi \int_0^{\infty} r g_r(r) J_k(2\pi r\rho)\, dr.$$

In these expressions ρ and ϕ are the coordinates in the transform space, and J_k is the Bessel function of the first kind of order k.

Applying this result to (22) yields

$$F[I_s] = \text{circ}\left(\frac{\rho R_o}{\gamma}\right) 2\pi \sum_{k=-\infty}^{\infty} c_k(-j)^k e^{jk\phi} \int_0^{R_o} r J_k(2\pi r\rho)\, dr \tag{23}^{\dagger}$$

where

$$c_k = \begin{cases} \frac{1}{2}(\exp[j\gamma_1] + \exp[j\gamma_2]), & k = 0 \\ 0 & k \text{ even}, \\ \frac{1}{\pi jk}(\exp[j\gamma_2] - \exp[j\gamma_1]), & k \text{ odd} \end{cases} \tag{24}$$

† Here we have used the easily established result that

$$F\left[\frac{\gamma}{rR_o} J_1(2\pi\gamma r/R_o)\right] = \text{circ}\,(\rho R_o/\gamma).$$

The appropriately normalized inverse transform of (23) is

$$\mathrm{I}_s = \int_0^\infty \int_0^{2\pi} F[\mathrm{I}_s]\rho \exp\left[j2\pi\tilde{r}_i\rho \cos(\theta_i - \phi)\right] d\rho\, d\phi$$

$$= 4\pi^2 \sum_{k=-\infty}^{\infty} c_k \exp\left[jk\theta_i\right] \int_0^{\gamma/R_o} \rho \int_0^{R_o} rJ_k(2\pi r\rho)\, drJ_k(2\pi\tilde{r}_i\rho)\, d\rho.$$

Here θ_i is the angular coordinate in the image plane, and we have already introduced $\tilde{r}_i = r_i/M$ for later simplicity. Now, letting $t = 2\pi R_o\rho$, $s = r/R_o$, and substituting expressions (24) for the coefficients c_k finally gives

$$\mathrm{I}_s = \frac{1}{2}(\exp[j\gamma_1] + \exp[j\gamma_2])h_o(r_i) + \frac{2}{\pi}(\exp[j\gamma_2] - \exp[j\gamma_1])$$

$$\cdot \sum_{k\ \mathrm{odd}} h_k(r_i) \frac{\sin k\theta_i}{k} \tag{25}$$

where

$$h_k(r_i) = \int_0^{2\pi\gamma} t \int_0^1 sJ_k(st)\, dsJ_k\left(\frac{r_i t}{MR_o}\right) dt. \tag{26}$$

Note that the integral $h_o(r_i)$ is just the integral given in (11), and that the integrals $h_k(r_i)$ are higher order generalizations of it.

The intensity in the image plane is obtained from equations (10) and (25) as

$$\bar{I}(r_i, \theta_i) = \begin{cases} |(ae^{i\alpha} - 1)\mathrm{I}_s + \exp[j\gamma_2]|^2, & 0 \leqq \theta_i \leqq \pi \\ & 0 \leqq r_i \leqq MR_o. \\ |(ae^{i\alpha} - 1)\mathrm{I}_s + \exp[j\gamma_1]|^2, & \pi < \theta_i < 2\pi \end{cases} \tag{27}$$

Clearly $\bar{I}$ depends upon both the radial and the angular coordinates in the image plane while the phase distribution over the entrance aperture depends only upon θ. Therefore, one would not expect, in general, that the semicircles would be visible in the intensity pattern. We will now demonstrate, however, that if a small value of $z = 2\pi\gamma$ is used for which the background is essentially uniform, and if $\gamma_2 - \gamma_1$ is small, the semicircles can be reproduced, and the jump in the normalized intensity will be proportional to $\gamma_2 - \gamma_1$, at least to first order. Moreover, as in our previous analysis, if $a = 1$, $\alpha = \pi/2$, and $z = 1.5$, the first condition will be satisfied, and the proportionality factor will be approximately unity.

To observe these facts we need only to substitute the first term of the

power series for the Bessel functions in (26) and conclude that for z sufficiently small

$$h_k(r_i) \approx \left(\frac{r_i}{MR_o}\right)^k \frac{z^{2k+2}}{(2^k k!)^2(k+2)(2k+2)}.$$

Thus for small z, the coefficients $h_k(r_i)$ are small, and they decrease extremely rapidly as k increases. Hence, if $\gamma_2 - \gamma_1$ is sufficiently small, the first term in (25) will dominate the second and I_s can be approximated by

$$I_s \approx \tfrac{1}{2}(\exp[j\gamma_1] + \exp[j\gamma_2])h_o(r_i).$$

With this approximation the normalized intensity is approximately

$$\tilde{I}(r_i, \theta_i) \approx \begin{cases} |\tfrac{1}{2}(ae^{j\alpha} - 1)(\exp[j\gamma_1] + \exp[j\gamma_2])h_o(r_i) + \exp[j\gamma_2]|^2, \\ \qquad\qquad 0 \leqq \theta_i \leqq \pi \\ \qquad\qquad 0 \leqq r_i \leqq MR_o. \\ |\tfrac{1}{2}(ae^{j\alpha} - 1)(\exp[j\gamma_1] + \exp[j\gamma_2])h_o(r_i) + \exp[j\gamma_1]|^2, \\ \qquad\qquad \pi < \theta_i < 2\pi. \end{cases}$$

Now, if the background is essentially uniform, the integral $h_o(r_i)$ in the above expression is essentially constant, and we observe that, for all practical purposes, $\tilde{I}$ is constant in each of the semicircular regions $0 \leqq \theta_i \leqq \pi$, $\pi < \theta_i < 2\pi$. Furthermore, the jump in $\tilde{I}$ between these two regions is just

$$\begin{aligned} \Delta\tilde{I} &\approx 2\ \mathrm{Re}\ \{\tfrac{1}{2}(ae^{j\alpha} - 1)(\exp[-j\gamma_2] - \exp[-j\gamma_1])(\exp[j\gamma_1] \\ &\quad + \exp[j\gamma_2])h_o(r_i)\} \\ &= 2a \sin\alpha\, h_o(r_i) \sin(\gamma_2 - \gamma_1) \\ &\approx 2a \sin\alpha\, h_o(r_i)(\gamma_2 - \gamma_1). \end{aligned}$$

From our previous work, we may recall that, for $z = 1.5$, $h_o(r_i) \approx 0.5$, and thus, for $a = 1$, $\alpha = \pi/2$, and $z = 1.5$,

$$\Delta\tilde{I} \approx (\gamma_2 - \gamma_1).$$

The same conditions that permitted radial discontinuities in the phase distribution to be qualitatively observed in the intensity pattern appear to permit a simple angular discontinuity such as (21) also to be observed. The representative results of numerical computations of the exact expressions (25), (26), (27) shown in Fig. 15 bear this out.

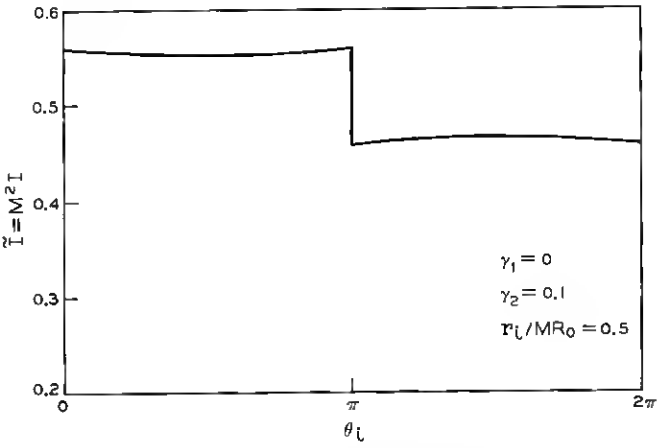

Fig. 15—Response to a semi-circular phase disk.

V. CONCLUSIONS

A general expression for the intensity distribution that is produced by a phase contrast imaging system with a circular phase plate was derived. It was shown that the results reduced to the well known expression for the phase contrast microscope when the size of the phase object is small compared to the area of the entrance aperture of the system. To obtain the intensity distribution for larger phase objects, a convolution integral must be evaluated by analytical and/or numerical techniques.

The intensity patterns that are produced by phase disks, phase rings, and semicircular phase distributions were derived, and the results of numerical computations that were based on these derivations were studied. In general, it appeared that qualitative reproductions of these simple phase distributions could be observed in the intensity patterns if the size of the phase spot on the phase plate was chosen to yield a uniform background, if the resulting parameter, z, was small, and if the magnitudes of the phase perturbations were small enough so that the approximation $e^{j\phi} \approx 1 + j\phi$ held. Moreover, it appeared that, if these conditions were not met, rather poor reproductions of even these simple phase distributions could result.

VI. ACKNOWLEDGMENTS

The author would like to thank J. Alan Cochran, E. R. Nagelberg, and N. J. Kolettis for helpful discussions on this research.

REFERENCES

1. Zernike, F., Z. Tech. Physik, *16* (1935), p. 454 ff.
2. Zernike, F., Physica *IX*, No. 7 (1942), p. 686 ff.
3. Zernike, F., Physica *IX*, No. 10 (1942), p. 974 ff.
4. Hopkins, H. H., Proc. Phys. Soc. B, *66* (1953), p. 331 ff.
5. Born, M., and Wolf, E., *Principles of Optics,* New York: Pergamon Press, 1965.
6. Hopkins, H. H., Rev. Opt., *31* (1952), p. 142 ff.
7. De, M., and Som, S. C., J. Opt. Soc. Amer., *53* (1963), p. 779 ff.
8. De, M., and Mondal, P. K., Optica Acta, *17* (1970), p. 397 ff.
9. Goodman, J. W., *Introduction to Fourier Optics,* New York: McGraw-Hill, 1959.
10. Mondal, P. K., and Slansky, S., Optica Acta, *17* (1970), p. 425 ff.